Lasers in Surgery and Medicine 19:46–62 (1996)

Development of a Multivariate Statistical Algorithm to Analyze Human Cervical Tissue Fluorescence Spectra Acquired In Vivo

Nirmala Ramanujam, PhD, Michele Follen Mitchell, MD, Anita Mahadevan, MS, Sharon Thomsen, MD, Anais Malpica, MD, Thomas Wright, MD, Neely Atkinson, PhD, and Rebecca Richards-Kortum, PhD

Biomedical Engineering Program, University of Texas, Austin, Texas 78705 (N.R., A.Mah., R.R.K.); Departments of Gynecology (M.F.M.), Pathology (S.T., A.Mal.), and Biomathematics (N.A.), UT MD Anderson Cancer Center, Houston, Texas, 77035; Department of Pathology, Columbia University, New York City, New York (T.W.)

Background and Objective: **A general multivariate statistical algorithm has been developed to analyze the diagnostic content of cervical tissue fluorescence spectra acquired in vivo.**
Materials and Methods: **The primary steps of the algorithm are to: (1) preprocess the data to reduce inter-patient and intra-patient variation of tissue spectra within a diagnostic category, without a priori information, (2) dimensionally reduce the preprocessed fluorescence emission spectrum with minimal information loss and use it to select the minimum number of the original emission variables of the fluorescence spectrum required to achieve classification with negligible decrease in predictive ability, and (3) assign a posterior probability to the diagnosis of each sample, so that samples with relative uncertain diagnosis can be reevaluated by a clinician. The algorithm was tested retrospectively and prospectively on cervical tissue spectra acquired from 476 sites from 92 patients at 337 nm excitation.**
Results: **The algorithm based on the entire fluorescence spectrum differentiates squamous intraepithelial lesions (SILs) from normal squamous epithelia and inflammation with an average sensitivity and specificity of 88%±1.4 and 70%±1, respectively. The average sensitivity and specificity of the identical algorithm based on intensity selected at only two emission wavelengths is 88%±1.4 and 71%±1.4, respectively.**
Conclusion: **The multivariate statistical algorithm based on both types of spectral inputs at 337 nm excitation has a similar sensitivity and significantly improved specificity relative to colposcopy in expert hands.**

Key words: cervix, in vivo diagnosis, laser-induced fluorescence, neoplasia, squamous intraepithelial lesion (SIL)

INTRODUCTION

Cervical cancer is the second most common malignancy in women worldwide, exceeded only by breast cancer [1]. In the United States, 15,000 new cases of invasive cervical cancer and 55,000 cases of carcinoma in situ (CIS) were reported in 1994 [2]. Women < 35 years old account for up to 24.5% of patients with invasive cervical cancer, and the incidence continues to increase for women in this age group [1].

The mortality associated with cervical cancer can be reduced if this disease is detected at the

Accepted for publication October 4, 1995.

Address reprint requests to Dr. Rebecca Richards-Kortum, Department of Electrical and Computer Engineering, University of Texas, ENS 610, Austin, TX 78705.

early stages of development or at the pre-cancerous state (Cervical Intraepithelial Neoplasia (CIN)). Currently, a Pap smear is used to screen for CIN and cervical cancer in the general female population. Although it is a relatively simple technique, the Pap smear has been reported to have a false-negative error rate of 15–40% [3]. Patients with an abnormal Pap smear are followed up by a diagnostic procedure called colposcopy. In a colposcopic examination, acetic acid is applied to cervix, which turns abnormal sites aceto-white [4]. The colposcope (low power microscope used to view the cervix with reflected white light) is then used to direct biopsies of these sites for histologic confirmation. Colposcopy requires extensive training and its accuracy for diagnosis is variable and limited. In expert hands [5], the average sensitivity and specificity of colposcopy for differentiating diseased tissues (squamous intraepithelial lesions (SILs)) from nondiseased tissues (normal squamous epithelia and inflammation) have been reported to be 94%±6 and 48%±23, respectively; its average sensitivity and specificity for differentiating high grade SILs (CIN II, CIN III, CIS) from low grade SILs (Human Papilloma Viral (HPV) infection, CIN I) are 79%±23 and 66%±18, respectively. A diagnostic method that could improve the accuracy of colposcopy in the hands of less experienced practitioners could allow more effective wide scale diagnosis and potentially reduce the mortality associated with cervical cancer.

A potential diagnostic technique for cervical cancer and its precursor is laser-induced fluorescence spectroscopy. There is much evidence to indicate that fluorescence spectroscopy can be used to identify neoplastic tissues in a variety of organ systems in vitro and in vivo [6–11]. Techniques based on fluorescence spectroscopy have an advantage over current detection methods for cervical diseases in that they are able to quantitatively detect changes in cellular chemistry and tissue architecture associated with progression of disease in a fast and nondestructive manner. Therefore, successful application of this technique will clinically represent an important step forward in medical laser spectroscopy and gynecologic oncology.

Previously, a two-stage algorithm developed from laser-induced fluorescence spectra of 115 cervical sites from 28 patients at 337 nm excitation demonstrated that fluorescence spectroscopy can be used to differentially diagnose SILs in vivo [12, 13]. A retrospective and prospective evaluation of the algorithm's performance using fluorescence spectra from 476 cervical sites in 92 patients [14] demonstrated that stage 1 of the algorithm can differentiate SILs HPV, CIN I, CIN II, and CIN III) from normal squamous mucosa and inflammation with an average sensitivity and specificity of 85%±1 and 76%±1, respectively. These results are similar to those reported in the clinical study in which the algorithm was initially developed [12,13] and represents a similar sensitivity and a significantly improved specificity relative to colposcopy in expert hands [5]. Stage 2 of the algorithm differentiated high grade SILs from low grade SILs with an average sensitivity and specificity of 80%±4 and 51%±11, respectively. The specificity of the algorithm is significantly lower (25% decrease) than its specificity in the clinical study in which it was initially developed [12, 13] and relative to that of colposcopy in expert hands.

There are several limitations of the diagnostic algorithm presented that must be addressed prior to clinical implementation. A first limitation is the algorithm as presented requires a fluorescence measurement from a colposcopically normal reference site in each patient in order to compensate for interpatient variation and then determine whether an unknown site is abnormal. Although this indicates that measurement of fluorescence from a normal site in each patient is needed to serve as a control, an alternate method that can still account for interpatient variation without the need to identify a reference normal site on the basis of colposcopy alone would be clinically more useful. A second limitation of the two-stage algorithm is that the method of data reduction used to determine diagnostically important parameters was not systematic; simple spectral parameters incorporating fluorescence intensity at several emission wavelengths were used as inputs to the diagnostic algorithm. Although intensity at a few emission wavelengths can be utilized to develop an effective algorithm, a thorough and systematic approach needs to be employed to determine if intensities at other emission wavelengths may provide additional information. A final limitation of the algorithm is that a simple decision line that minimizes misclassification was used to discriminate between classes of interest. This binary classification scheme provides no information about the certainty of an assigned diagnosis and is likely to exhibit poor performance near boundaries between classes. Therefore, a probability associated with classifying a cervical

sample into a particular diagnostic category would be more useful to the clinician.

The primary goal of this study was to develop a multivariate statistical algorithm to overcome the limitations of the two-stage algorithm. Preprocessing methods that do not require a priori information were utilized to calibrate for interpatient and intrapatient variation in tissue fluorescence spectra. Principal Component Analysis (PCA) was used to dimensionally reduce the spectral data with minimal information loss and, finally, a probability based classification algorithm based on logistic discrimination was developed using these principal components. The diagnostic algorithm was developed and tested retrospectively and prospectively using cervical tissue fluorescence spectra acquired at 337 nm excitation from 476 sites in 92 patients. The performance of the new algorithm was then compared to that of colposcopy in expert hands [5] to evaluate the clinical value and diagnostic limitations of this algorithm. The multivariate statistical algorithm presented here is general and can be applied to spectral data acquired at any excitation wavelength.

A second goal of our work was to determine if the contribution of tissue spectra at multiple excitation wavelengths can enhance the performance of a diagnostic algorithm at 337 nm excitation. To address this objective, tissue spectra were acquired at additional excitation wavelengths: 380 and 460 nm, from a subset of the 92 patients examined. The selection of these excitation wavelengths was based on the results of an in vitro study that demonstrated that differences in fluorescence spectra of histologically diseased and nondiseased cervical tissues are greatest near 340, 380, and 460 nm, with maximal differences at 340 nm excitation [11]. In a companion report [15], tissue spectra acquired in vivo at 380 and 460 nm excitation are analyzed using the multivariate statistical algorithm developed using the larger spectral data set at 337 nm excitation here. The companion work also presents a contrast of the diagnostic content of spectra at these excitation wavelengths to that at 337 nm excitation, to determine optimal excitation wavelengths for the differential diagnosis of SILs in vivo.

MATERIALS AND METHODS

Instrumentation

A spectroscopic system incorporating a pulsed nitrogen pumped-dye laser, an optical fiber probe, and an optical multichannel analzyer was utilized to record fluorescence spectra from the intact cervix at colposcopy (Fig. 1a). The common end of the probe consists of a central fiber surrounded by a circular array of six fibers. All seven fibers have the same characteristics (0.22 NA, 200 μm core diameter). Two of the peripheral fibers (Fig. 1b) deliver excitation light to the tissue surface; one delivers excitation light from the nitrogen laser and the other, from the dye module (overlap of the illumination area viewed by both optical fibers is $> 85\%$). The purpose of the remaining five fibers is to collect the emitted fluorescence from the tissue surface directly illuminated by each excitation fiber (overlap of the illumination and collection site is $\sim 85\%$). A quartz shield is placed at the tip of the probe (Fig. 1c) to provide a fixed distance between the fibers and the tissue surface, so fluorescence intensity can be reported in calibrated units.

Excitation light at 337 nm excitation was focused into the proximal end of one of the two excitation fibers to produce a 1 mm diameter spot at the outer face of the shield. Excitation light from the dye module coupled into the second excitation fiber was produced by using appropriate fluorescent dyes; BBQ (1E-03M in seven parts toluene and three parts ethanol) was used to generate light at 380 nm excitation and Coumarin 460 (1E-02 M in ethanol) was used to generate light at 460 nm excitation. The average transmitted pulse energies at 337, 380, and 460 nm excitation were 20, 12, and 25 μJ, respectively. The laser characteristics for this study were a 5 ns pulse duration and a repetition rate of 30 Hz.

The proximal ends of the collection fibers were arranged in a circular array and imaged at the entrance slit of a polychromator (Jarrell Ash, Monospec 18) coupled to an intensified 1,024-diode array controlled by a multichannel analyzer (Princeton Instruments, OMA, Princeton, NJ). 370, 400 and 470 nm long pass filters were used to block scattered excitation light at 337, 380 and 460 nm excitation, respectively. A 205 ns collection gate, synchronized to the leading edge of the laser pulse using a pulser (Princeton Instruments, PG200), effectively eliminated the effects of the colposcope's white light illumination during fluorescence measurements. Data acquisition was computer controlled.

Laser-Induced Fluorescence Spectroscopy at Colposcopy

The patients examined with this spectroscopic system were selected from a group of pa-

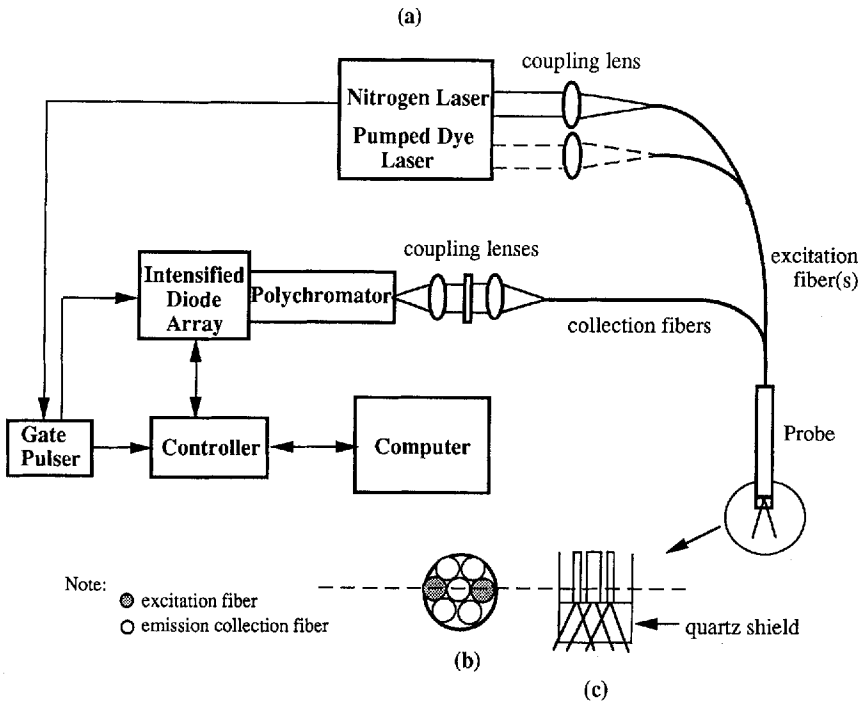

Fig. 1. Spectroscopic system used to measure fluorescence spectra from the intact cervix at two excitation wavelengths.

tients referred for colposcopy for suspected CIN on the basis of an abnormal Pap smear. After colposcopic examination but before biopsy, the rigid end of the probe was inserted into the vagina under colposcopic control, until the tip was in contact with the cervical tissue surface and then fluorescence spectra were acquired. On average, spectra from two abnormal sites and two normal sites were obtained from each patient. Spectra were acquired from normal sites on the squamous epithelium of the ectocervix. In some patients, columnar epithelia contained within the endocervical canal extended out through the external os onto the colposcopically visible portion of the cervix. If any columnar epithelium was visible in a patient, a normal spectrum was acquired from this region as well. In this study, normal columnar epithelia were visible in ~ 50% of the patients examined spectroscopically. Spectra were acquired from abnormal sites on the squamous epithelium or from areas surrounding the squamocolumnar junction.

Prior to each patient study, the probe was disinfected for 20 minutes in Metricide (Metrex Corp.). A background spectrum was then measured at 337 nm excitation and then at either 380 or 460 nm excitation, with the probe dipped in a nonfluorescent bottle containing distilled water. This background spectrum was subtracted from all subsequently acquired data at corresponding excitation wavelengths for that patient. Next, five spectra were measured from a Rhodamine calibration standard at both excitation wavelengths. After calibration, fluorescence spectra were acquired from the cervix: 15 spectra for 15 consecutive laser pulses were acquired at 337 nm excitation, and then 100 spectra for 100 consecutive laser pulses were acquired at either 380 or 460 nm excitation wavelength (generated using the appropriate dye) from each cervical site. Spectra were collected in the visible region of the electromagnetic spectrum with a resolution of 10 nm (full width at half maximum) and a signal-to-noise ratio of 30:1 at the fluorescence maximum at each excitation wavelength. No appreciable photobleaching was observed between the collection of the initial and final spectra of the Rhodamine calibration standard or the cervical sites.

All spectra were corrected for the nonuniform spectral response of the detection system using correction factors obtained by recording the spectrum of an N.I.S.T. (National Institute of Standards and Technology) traceable calibration

tungsten ribbon filament lamp. Then in each patient, the average spectrum from each cervical site was divided by the average peak fluorescence intensity of the Rhodamine calibration standard, at each excitation wavelength; absolute fluorescence intensities are reported in these calibrated units.

Fluorescence spectra were acquired from a total of 476 sites in 92 patients at 337 nm excitation. In 40 of the 92 patients, fluorescence spectra were acquired from 165 sites at both 337 and 380 nm excitation. In 24 of the remaining 52 patients, fluorescence spectra were acquired from 147 sites at both 337 and 460 nm excitation. Analysis of spectra at 337 nm excitation are presented here; analysis of spectra at 380 and 460 nm excitation are presented in a companion work [15].

Pathology

Biopsies were taken only from the abnormal sites examined spectroscopically by the probe (± 1 mm) to comply with routine patient care procedure. The biopsies were fixed in formalin and submitted for routine histologic examination. Hemotoxylin and eosin stained sections were evaluated independently by three board certified pathologists. Discrepant samples were reevaluated to arrive at a consensus diagnosis during a conference specifically designed to set criteria for reading slides.

All specimens were classified using the Bethesda system, a new classification scheme that reflects current clinical managment of cervical disease [1]. The specimens were classified as normal squamous or normal columnar mucosa, inflammation, low grade SIL (HPV, CIN I) or high grade SIL (CIN II, CIN III, CIS). Samples with multiple diagnoses were classified according to the most severe diagnosis.

Clinically, tissues with inflammation are considered nondiseased and do not require treatment. Low grade and high grade SILs are considered to be diseased. However, only selected patients with low grade SIL are generally followed up, whereas all patients with high grade SIL are treated [1]. Therefore, the goal of a clinically useful diagnostic algorithm should be to: (1) differentiate diseased (low grade SIL, high grade SIL) from non diseased tissues (normal epithelia, inflammation), and (2) differentiate high grade SILs from low grade SILs. In the current study, the algorithm was developed and optimized to reflect this classification scheme.

Multivariate Statistical Algorithm Development

The six primary steps involved in the development of the multivariate statistical algorithm are: (1) preprocessing of spectral data to account for interpatient and intrapatient variation of spectra from a diagnostic category, (2) partitioning of the preprocessed spectral data from all patients into calibration and prediction data sets, (3) dimension reduction of the preprocessed spectra from the calibration set using PCA, (4) selection of the diagnostically most useful principal components using a one-sided unpaired student's t-test, (5) development of a probability-based classification scheme based on logistic discrimination using the diagnostically useful principal component scores of the calibration set as inputs, and (6) a retrospective evaluation of the algorithm's performance on the calibration data set and a prospective evaluation of the algorithm's performance on the prediction data set.

Preprocessing. The objective of preprocessing is to calibrate tissue spectra for interpatient and intrapatient variation that might obscure differences in the spectra of different tissue types. Three methods of preprocessing (described in detail below) were invoked on the spectral data: (1) normalization, (2) mean-scaling, and (3) a combination of normalization and mean-scaling. All three methods of preprocessing were compared to determine which provides the greatest diagnostic information.

Each spectrum was normalized by dividing the fluorescence intensity at each emission wavelength by the maximum fluorescence intensity of that sample. Normalizing a fluorescence spectrum removes absolute intensity information; algorithms developed from normalized fluorescence spectra rely on differences in spectral line shape information for diagnosis. If the contribution of the absolute intensity information is not significant and there is negligible interpatient variation in spectral line shape, two advantages are realized by utilizing normalized spectra: (1) it is no longer necessary to calibrate for interpatient or intrapatient variation of normal tissue fluorescence intensity, and (2) fluorescence intensity does not need to be recorded in calibrated units.

Mean-scaling was performed by calculating the mean spectrum for a patient (using all spectra obtained from cervical sites in that patient) and subtracting it from each spectrum in that patient. Mean-scaling can be performed on both unnormalized (original) and normalized spectra. Mean-

TABLE 1. Histopathologic Classification of Sample Spectra at 337 nm excitation from Calibration and Prediction Data Sets

Classification	Calibration set	Prediction set
Normal squamous	127	126
Normal columnar	25	25
Inflammation	16	16
Low grade SIL	40	40
High grade SIL	31	30

scaling does not require colposcopy to identify a reference normal site in each patient prior to spectroscopic analysis. However, unlike normalization, mean-scaling displays the differences in the fluorescence spectrum from a particular site with respect to the average spectrum from that patient. Therefore, this method maximally enhances differences in fluorescence spectra between tissue categories when spectra are acquired from approximately equal numbers of nondiseased and diseased sites from each patient.

Development of calibration and prediction data sets. Fluorescence spectra from the data set at 337 nm excitation were randomly assigned to either a calibration set or prediction set with the condition that both sets contain roughly equal number of samples from each histopathologic category (Table 1). The random assignment ensured that not all spectra from a single patient were contained in the same data set. The purpose of the calibration set was to develop and optimize the algorithm, and the purpose of the prediction set was to prospectively test its accuracy in an unbiased manner.

Principal Component Analysis. PCA was used to transform the original variables of the fluorescence emission spectra into a smaller set of linear combinations of the original variables called principal components that account for most of the variance of the original data set. Although PCA [16] may not provide direct insight to the morphologic and biochemical basis of tissue spectra, it provides a novel approach of condensing all the spectral information into a few manageable components, with minimal information loss. Furthermore, each principal component can be easily related to the original emission spectrum, thus providing insight into diagnostically useful emission variables.

Prior to PCA, a data matrix is created where each row of the matrix contains the preprocessed fluorescence spectrum of a sample and each column contains the preprocessed fluorescence intensity at each emission wavelength. The data matrix D (r x c), consisting of r rows (corresponding to total r samples) and c columns (corresponding to intensity at c emission wavelengths) can be written as:

$$D = \begin{pmatrix} D_{11}\ D_{12} & .....D_{1c} \\ D_{21}\ D_{22} & .....D_{2c} \\ . & \\ . & \\ D_{r1}\ D_{r2} & .....D_{rc} \end{pmatrix} \quad (1)$$

The first step in PCA is to calculate the covariance matrix, Z (eq.(2)), where D_m represents the mean-sided data matrix.

$$Z = \frac{1}{r-1}(D_m{}'D_m) \quad (2)$$

The square covariance matrix, Z (c x c), is decomposed into its respective eigenvalues (λ_j) and eigenvectors. Because of experimental error, the total number of eigenvalues will always equal the total number of columns (c) in the data matrix D assuming that $c < r$. The goal is to select $n < c$ eigenvalues that can describe most of the variance of the original data matrix to within experimental error. The variance, V, accounted for by the first n eigenvalues can be calculated using eq.(3), where λ_j is the jth eigenvalue:

$$V = 100\left(\frac{\sum_{j=1}^{n}\lambda_j}{\sum_{j=1}^{c}\lambda_j}\right) \quad (3)$$

Next, the principal component score matrix can be calculated according to eq. (4):

$$R = DC \quad (4)$$

where, D (r x c), is the preprocessed data matrix and C (c x n) is a matrix whose columns contain the n eigenvectors that correspond to the first n eigenvalues. Each row of the principal component score matrix, R (r x n), corresponds to the principal component scores of a sample and each column corresponds to a principal component. The principal components are mutually orthogonal.

Finally, the component loading is calculated for each principal component. The component loading represents the correlation between the principal component and the variables of the original fluorescence emission spectrum. The component loading can be calculated using eq. (5):

$$CL_{ij} = \frac{C_{ij}}{\sqrt{S_{ii}}}\sqrt{\lambda_j} \quad (5)$$

where CL_{ij} represents the correlation between the ith variable (preprocessed intensity at ith emission wavelength) and the jth principal component. C_{ij} is the ith component of the jth eigenvector, λ_j is the jth eigenvalue, and S_{ii} is the variance of the ith variable.

PCA was performed on each type of preprocessed data matrix, described above; n eigenvalues accounting for 99% of the variance of the original preprocessed data set were retained (eq. 3). The corresponding n eigenvectors were then multiplied by the original data matrix to obtain the principal component score matrix R. Finally, the component loading of each principal component was calculated.

Student's t-test. Average values of principal component scores were calculated for each principal component for each diagnostic category obtained from the preprocessed data matrix. A one-sided unpaired student's t-test [17] was employed to determine the diagnostic contribution of each principal component. The hypothesis that the means of the principal component scores of two diagnostic categories are different was tested for: (1) normal squamous epithelia and SILs, (2) normal columnar epithelia and SILs, and (3) inflammation and SILs. The t-test was extended a step further to determine if there are any statistically significant differences between the means of the principal component scores of high grade SILs and low grade SILs. Principal components for which the hypotheses stated above was true below the 0.05 level of significance were retained for further analysis.

Logistic discrimination. Logistic discrimination [18] was used to develop classification based on posterior probabilities, overcoming the drawback of a binary decision scheme. This statistical classification algorithm method is based on Bayes theorem and can be used to calculate the posterior probability that an unknown sample belongs to each of the possible diagnostic categories identified. Classifying the unknown sample into the diagnostic category for which its posterior probability is highest results in a classification scheme that minimizes the rate of misclassification.

For two diagnostic categories, G_1 and G_2, the posterior probability of being a member of G_1, given measurement, x, is:

$$P(G_1|x) = \frac{P(x|G_1)P(G_1)C(2|1)}{P(x|G_1)P(G_1)C(2|1) + P(x|G_2)P(G_2)C(1|2)} \quad (6)$$

where $P(x|G_i)$ is the conditional probability that a tissue sample of type i will have principal component score x, and $P(G_i)$ is the prior probability of finding tissue type i in the sample population. The prior probability $P(G_i)$ is an estimate of the likelihood that a sample of type i belongs to a particular group when no information about it is available. If the sample size is considered representative of the population, the observed proportions of cases in each group can serve as estimates of the prior probabilities. C(j|i) is the cost of misclassifying a sample into group j when the actual membership is group i. The cost of misclassification of a diagnostic category can be varied from 0 to 1 with the condition that the sum of the cost of misclassification for all diagnostic categories under consideration equals to one.

The conditional probabilities can be obtained from the probability distributions of the n principal component scores for each tissue type i. The probability distributions can be modeled using the gamma function [16], which is characterized by two parameters, α and β. The best fit of the gamma function to the probability distribution of each principal component (score) for each tissue type can be obtained in the least squares sense, using as α and β as free parameters of the fit. The gamma function can then be used to calculate the conditional probability that an unknown sample, given that it is from tissue type i, will exhibit the principal component score x. If more than one principal component is needed to describe a sample population, then the conditional joint probability is simply the product of the conditional probability of each principal component (assuming that each principal component is an independent variable) for that sample population.

RESULTS

Typical Spectra

Figure 2a,b illustrates average spectra per site acquired from all cervical sites in two typical

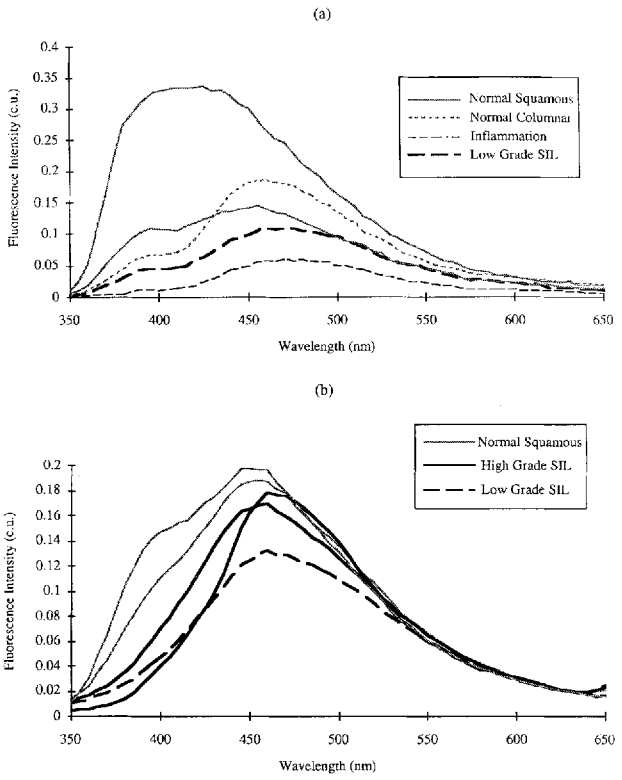

Fig. 2. Fluorescence spectra from two typical patients at 337 nm excitation.

patients at 337 nm excitation. All fluorescence intensities are reported in the same set of calibrated units. Figure 2a,b indicates that the fluorescence intensities of normal squamous epithelia can vary by a factor of two from patient to patient; within a single patient, the variation is generally $< 15\%$ (Fig. 2b), but in some patients, the variation of normal tissue fluorescence intensity can be as much as 50% (Fig. 2a). Although the fluorescence intensities of SILs are lower than those of corresponding normal squamous epithelia within a patient in general (Fig. 2a), in some instances this is not as dramatic, as illustrated in Figure 2b, where the fluorescence maxima of the high grade SILs are within 10% that of corresponding normal squamous epithelia. Examination of Figure 2a indicates that the fluorescence intensities of the normal columnar sample and sample with inflammation are similar to that of SILs.

The spectral line shape and peak emission wavelength of normal squamous tissues display a large variation from patient to patient; within a patient, the variation is not as significant. The peak emission wavelength of normal squamous tissue spectra range from 430 to 460 nm, whereas that of SILs range from 450 to 480 nm. The differences in peak emission wavelength between the spectra of normal squamous epithelia and SILs imply differences in their spectral line shape. Note that the peak emission wavelength of the spectrum of the normal columnar sample and that of the sample with inflammation lie in the same range as that of the SIL (Fig. 2a).

Results of Multivariate Statistical Algorithm

Preprocessing. Spectra from each patient at 337 nm excitation were evaluated using three methods of preprocessing: (1) normalization, (2) mean-scaling, and (3) a combination of normalization and mean-scaling. Tissue spectra processed by all three methods were evaluated using the multivariate statistical algorithm. The algorithm based on all three methods demonstrated significant intercategory differences between SILs and normal squamous epithelia only. The algorithm based on spectra that were normalized and then mean-scaled demonstrated the greatest differences between normal squamous tissues and SILs. Figure 3a–f illustrates and compares the three types of corresponding preprocessed spectra of the two typical patient spectra in Figure 2a,b.

Figure 3a,d shows the preprocessed spectra resulting from normalization of the typical spectra shown in (a) and (b) respectively, in Figure 2. Following normalization, the fluorescence intensities of normal squamous tissues are greater than those of the corresponding SILs over the wavelength range 380 to 450 nm. From 460 to 580 nm, the fluorescence intensities of SILs are greater than those of normal squamous tissues, which in part reflects the longer peak emission wavelength of SILs (Fig. 3d). Spectral line shapes and emission maximas of normal columnar samples and tissues with inflammation are very similar to that of the low grade SIL as shown in Fig. 3a. Preprocessing by normalization clearly reduces the interpatient and intrapatient variation in fluorescence intensity. However, patient-to-patient variation in fluorescence line shape of normal squamous tissues is still great as evidenced in Figure 3a,d.

Figure 3b,e shows the corresponding mean-scaled spectra of the original spectra shown in (a) and (b), respectively, in Figure 2. The fluorescence of normal squamous tissues are in general

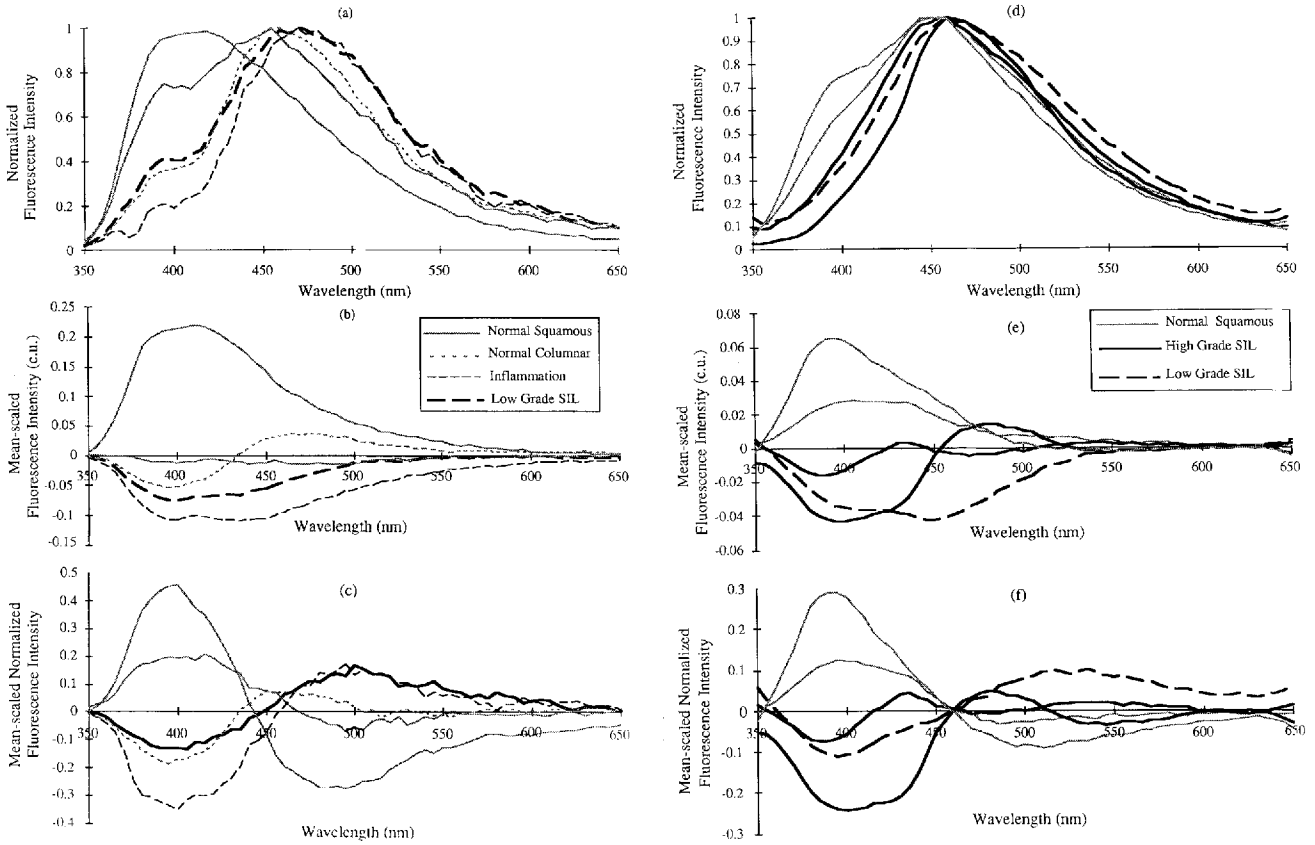

Fig. 3. (**a**) and (**d**), normalized spectra; (**b**) and (**e**), mean-scaled spectra; (**c**) and (**f**) normalized mean-scaled spectra of the two typical patient spectra at 337 nm excitation shown in Figure 2.

greater than the mean, whereas that of SILs are less than the mean over almost the entire emission spectrum. There is considerably less patient-to-patient variation in the fluorescence spectra of normal squamous tissues and SILs when data are preprocessed by mean-scaling. However, in patients with a large variation in the fluorescence intensity of normal squamous tissues (Fig. 2a), mean-scaling unfortunately preserves this intrapatient variation (Fig. 3b).

Figure 3c,f shows the corresponding spectra preprocessed with the final method used here: normalization, followed by mean-scaling. Here, the fluorescence line shapes of normal squamous tissues display much less patient-to-patient variation compared to spectra in Figure 3a,d. Further, the intrapatient variation in normal squamous tissue fluorescence line shape is significantly smaller (Fig. 3c) than the intrapatient variation in normal squamous tissue fluorescence intensity (Fig. 3b). Therefore, preprocessing by first normalizing all spectra from a patient to a peak intensity of unity, followed by mean-scaling all spectra within a patient provides the greatest reduction in intra- and interpatient variation in the spectra of tissues from a particular diagnostic category. For this reason, all subsequent data analysis was carried out using normalized, mean-scaled spectral data.

Principal Component Analysis. PCA of normalized, mean-scaled spectra at 337 nm excitation from the calibration data set resulted in three principal components accounting for 99% of the total variance. Only the first two principal components demonstrated the statistically most significant differences ($P < 0.05$) between normal squamous tissues and SILs (PC1: p < 1E-25, PC2: $P < 0.006$). The one-tail P values of the scores of the third principal component were not statistically significant ($P > 0.2$). Therefore, further analysis was performed using these two principal components.

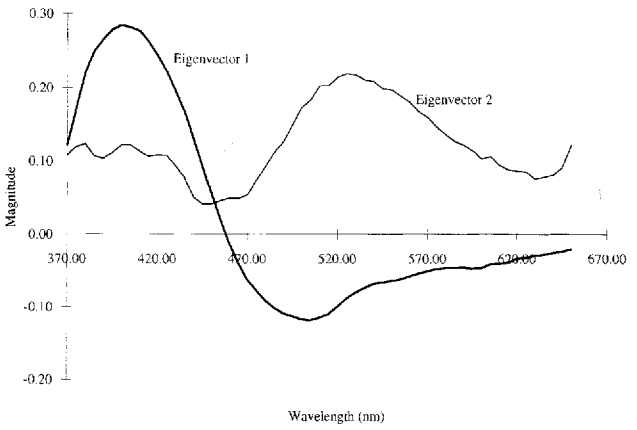

Fig. 4. Two eigenvectors retained from the principal component analysis of the preprocessed calibration data matrix containing normalized, mean-scaled spectra at 337 nm excitation.

Figure 4 displays the eigenvectors from which the first two principal components were calculated. Eigenvector 1 corresponds to the largest eigenvalue, which accounts for 85% of the total variance. Eigenvector 2 corresponds to the second largest eigenvalue, which accounts for 10% of the total variance. Evaluation of Figure 4 illustrates that eigenvector 1 has a very similar line shape to that of normalized, mean-scaled spectra of normal squamous tissues (Fig. 3c,f). The magnitude of the vector is greater than zero with a maximum near 400 nm, for wavelengths below 450 nm; above 450 nm, its magnitude is less than zero, with a negative maximum at 500 nm. Inversion of this eigenvector, about the line $Y = 0$, results in a spectral line shape similar to that of SILs. The second eigenvector describes spectral line shape primarily over the wavelength range, 480–600 nm, with a maximum at 540 nm.

A scattergram of principal component scores 1 (PC1) and 2 (PC2) is shown in Figure 5 for normal squamous samples, normal columnar samples, samples with inflammation, and SILs from the calibration data set. Evaluation of Figure 5 indicates that the mean of PC1 (3.38) of normal squamous epithelia is greater than that of SILs (2.66), whereas the variance of the two tissues types are similar (normal squamous: 0.46, SILs: 0.35). The opposite phenomenon is observed for PC2. The mean (1.00) of PC2 is similar for both tissue types, whereas the variance of PC2 of SILs (0.24) is approximately twice that of normal squamous tissues (0.14). The correlation between PC1 and PC2 is zero. Note that the principal component scores of normal columnar epithelia and inflammation are indistinguishable from that of SILs. Therefore, the principal component scores of these samples were not included in the development of the classification algorithm.

Logistic discrimination. Figure 6a,b shows a comparison of the measured probability distribution and best fit gamma function of PC 1 and PC 2 of normal squamous tissues and SILs, respectively, from the calibration data set. In all instances, there is excellent agreement between the experimentally measured and calculated gamma distributions. The gamma function was used to calculate the conditional joint probabilities of PC 1 and PC 2 for normal squamous tissues and SILs. The prior probability was determined by calculating the percentage of each tissue type in the calibration set: 65% normal squamous tissues and 35% SILs. More generally, prior probabilities should be selected to describe the patient population under study; the values used here are appropriate as they describe the prediction set as well. The cost of misclassification was varied from 0 to 1 in 0.05 increments, and the optimal cost was identified when the total number of misclassified samples was a minimum. If the total number of misclassified samples was a minimum at more than one cost, then the cost at which the number of SILs misclassified was a minimum was selected. The optimal cost of misclassification for SILs was selected to be 0.55.

Posterior probabilities of belonging to each tissue type (normal squamous or SIL) were calculated for all samples in the calibration data set, using the known prior probabilities, cost of misclassification of SIL, and the conditional joint probabilities calculated from the gamma function. Figure 7 illustrates the posterior probability of belonging to the SIL category. The posterior probability is plotted for all samples in the calibration data set. 92% of SILs have a posterior probability > 0.5, and 80% of normal squamous tissues have a posterior probability < 0.5. A careful examination of the plot indicates that $> 80\%$ of the high grade SILs, and $> 85\%$ of low grade SILs have a posterior probability > 0.75. More than 70% of normal squamous tissues have a posterior probability < 0.25. Note that evaluation of normal columnar epithelia and samples with inflammation using this algorithm results in classifying a majority of them as SILs.

Figure 8 shows the percentage of normal squamous tissues and SILs correctly classified vs.

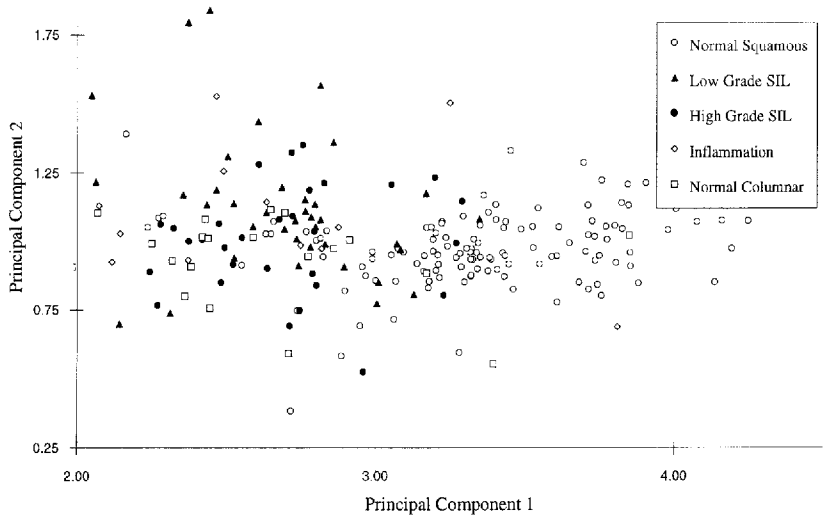

Fig. 5. Two-dimensional scattergram of the scores of principal component 1 (PC1) and principal component 2 (PC2) of normal squamous tissues, normal columnar tissues, samples with inflammation, and SILs from the calibration data set.

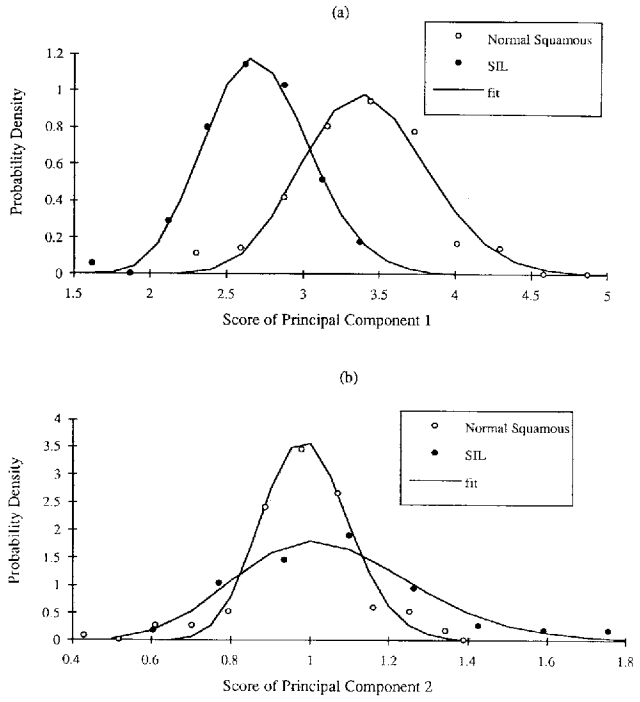

Fig. 6. Measured probability distributions and calculated gamma distributions of (a) principal component 1 scores (PC1) and (b) principal component 2 scores (PC2) of normal squamous tissues and SILs from the calibration data set.

cost of misclassification of SILs for the data from the calibration set. As expected, an increase in the SIL misclassification cost results in an increase in the proportion of correctly classified SILs and a decrease in the proportion of correctly classified normal squamous tissues. Note that varying the cost from .4 to .6 alters the correct classification of both SILs and normal squamous tissues by $< 15\%$, indicating that a small change in the cost does not significantly alter the performance of the algorithm for the prior probabilities used and is not sensitive to small changes in prior probability.

The algorithm was implemented on normalized, mean-scaled spectra of the prediction data set to obtain an unbiased estimate of its accuracy. The two eigenvectors obtained from the calibration data set were multiplied by the prediction data matrix to obtain the prediction set principal component score matrix. Using the same prior probabilities and cost of misclassification, and conditional joint probabilities calculated from the gamma function, all developed from the calibration data set, logistic discrimination was used to calculate the posterior probabilities for all samples in the prediction data set.

Confusion matrices in (a) and (b) in Table 2 show the spectroscopic classification using this algorithm on the calibration set and the prediction set, respectively. A comparison of the sample classification between the prediction and calibration sets indicates that there is only a 3% decrease in the proportion of correctly classified SILs and a 4% decrease in the proportion of correctly classi-

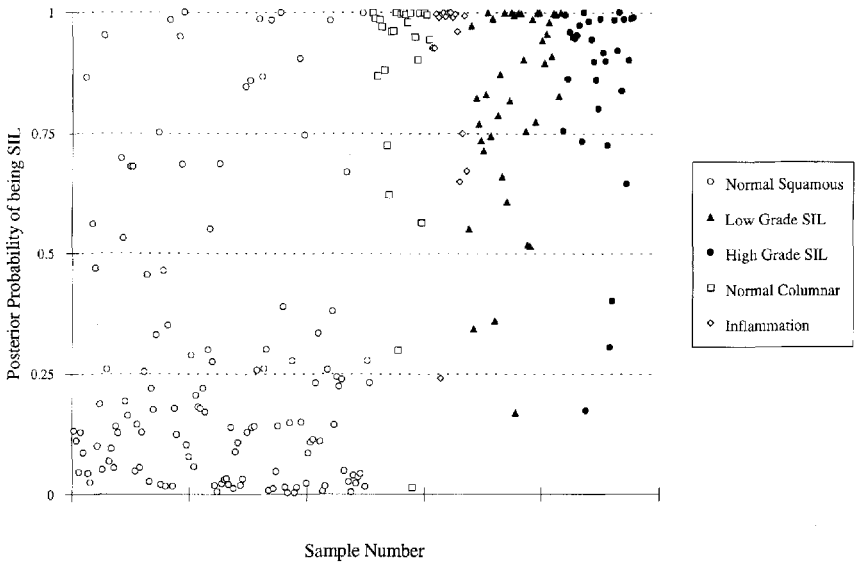

Fig. 7. The posterior probability of belonging to the SIL category of all samples in the calibration data set.

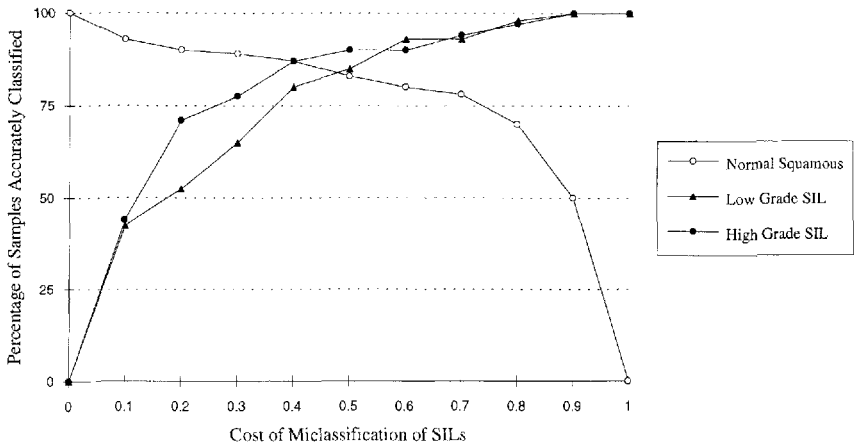

Fig. 8. Percentage of accurately classified normal squamous tissues and SILs vs. cost of misclassification of SILs for all SILs and normal squamous epithelia from the calibration data set.

fied normal squamous tissues on an unknown data set of approximately equal prior probability. This indicates that there is < 3% increase in the proportion of misclassified samples in the prediction data set relative to the calibration data set.

Analysis of fluorescence intensity at two emission wavelengths using the multivariate statistical algorithm. Figure 9 illustrates the component loadings of PC1 and PC2. The component loading represents the correlation between each principal component and the original preprocessed emission spectra of the calibration data set. The component loading can be interpreted by comparing it to the normalized mean-scaled spectra illustrated in Figure 3f. Component loading 1 is highly and positively correlated with the original emission spectrum over the wavelength range, 380–450, with a maximum at 400 nm (correlation = 1). Evaluation of the spectra in Figure 3f over this wavelength range indicates that spectra of normal squamous tissues are greater than the mean, whereas that of SILs are less than the mean; greatest differences between spectra of the two tissue types occurs at 400 nm, corresponding to the highest positive correlation in Figure 9. At 460 nm, the correlation of component loading 1 is near zero (Fig. 9); this corresponds to the intersection of spectra of normal squamous tissues and SILs at 460 nm (Fig. 3f). Above 460 nm, the correlation of component loading 1 decreases and

TABLE 2. Confusion Matrices Displaying Classification of Normal Squamous Tissues and SILS Using Multivariate Statistical Algorithm*

	(a)		
Classification	Normal Squamous	Low grade SIL	High grade SIL
Normal squamous	80%	7%	10%
SIL	20%	93%	90%
	(b)		
Classification	Normal squamous	Low grade SIL	High grade SIL
Normal squamous	76%	15%	7%
SIL	24%	85%	93%
	(c)		
Classification	Normal squamous	Low grade SIL	High grade SIL
Normal squamous	83%	16%	16%
SIL	17%	85%	84%
	(d)		
Classification	Normal squamous	Low grade SIL	High grade SIL
Normal squamous	87%	25%	10%
SIL	13%	75%	90%

*(**a**) and (**b**) display the classification accuracy of the algorithm applied to the *fluorescence emission spectra* of SILs and normal squamous tissues from the calibration and prediction sets, respectively. (**c**) and (**d**) contain the results of the algorithm applied to spectral intensities at two emission wavelengths (410 nm and 490 nm) of normal squamous tissues and SILs from the calibration and prediction sets, respectively.

reaches a negative maximum at 500 nm (correlation = −0.9). The negative correlation decreases to −0.7 at 520 nm and remains constant up to 600 nm. Evaluation of Figure 3f indicates that above 460 nm, the normalized, mean-scaled spectra of SILs are greater than the mean, whereas that of normal squamous tissues are less than the mean, reflecting the negative correlation; the maximum differences between these two tissue types for wavelengths > 460 nm occurs over the wavelength range 500–520 nm. Above 520 nm, the differences between the two tissue types remain fairly constant up to 600 nm. PC2 is much less correlated with the original emission spectrum with an average correlation of 0.5, but over the wavelength range, 520–570 nm, it is highly correlated with the emission spectrum with a maximum correlation of 0.7 at 540 nm.

We explored the utility of the component loadings for reducing the number of emission variables required to achieve classification with minimal decrease in predictive ability. Portions of the emission spectrum most highly correlated (correlation > 0.9 or < −0.9) with the component loadings were selected and the reduced data matrix was used to regenerate and evaluate the algorithm. Intensity at one emission wavelength that had the highest positive correlation over the wavelength range 350–450 nm and intensity at one emission wavelength that had the highest negative correlation over the wavelength range

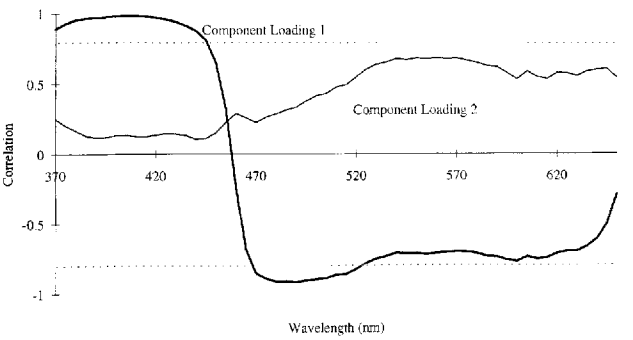

Fig. 9. Component loading for principal component 1 (PC1) and principal component 2 (PC2) from the calibration data set.

480–520 nm were selected. These variables were preprocessed fluorescence intensities at 410 nm and 490 nm. The algorithm was developed in an identical manner as was done with the entire fluorescence emission spectrum. It was optimized using the calibration set and implemented on the prediction data set.

Confusion matrices in (c) and (d) in Table 2 contain the results of the algorithm on the calibration and prediction data sets. The optimal cost of misclassification of SIL was selected to be 0.55. A comparison of the sample classification between the calibration and prediction data sets indicates that there is a 6% decrease in the proportion of correctly classified SILs, but a 4% increase in the proportion of correctly classified normal

TABLE 3. Comparison of Sensitivity and Specificity of Multivariate Statistical Algorithm Based on Entire Fluorescence Emission Spectrum and Intensity at Two Emission Wavelengths to Sensitivity and Specificity of Colposcopy in Expert Hands*

	Algorithm		Colposcopy	
Classification	Sensitivity	Specificity	Sensitivity	Specificity
Algorithm based on entire fluorescence emission spectrum	88% ± 1.4	70% ± 1.4	94% ± 6	48% ± 23
Algorithm based on fluorescence intensity at two emission wavelengths	88% ± 0	71% ± 0	94% ± 6	48% ± 23

*See ref. [5].

squamous tissues in the prediction data set. Overall, there was < 1% increase in the proportion of misclassified samples in the prediction data set relative to the calibration data set.

The sample classification in the calibration and prediction data sets based on the algorithm using the entire emission spectrum was compared to that using intensity at two emission wavelengths. Using the entire emission spectrum, the algorithm correctly classifies 91%±2 (mean and standard deviation) of SILs and 78%±3 of normal squamous tissues; using intensity at two emission wavelengths, the algorithm correctly classifies 82%±4 of SILs and 85%±3 of normal squamous epithelia. Although a lower percentage of SILs and a greater percentage normal squamous tissues are correctly classified using the latter algorithm, the total number of samples misclassified using both types of inputs are approximately the same.

Table 3 compares the performance of the multivariate statistical algorithm based on the entire emission spectrum and intensity at two emission wavelengths to that of colposcopy in expert hands for differentiating diseased (SILs) from nondiseased tissues (normal squamous epithelia and samples with inflammation) [5]. Normal columnar tissues were not included because the colposcopy data was obtained only from the ectocervix. The cost of misclassification of SILs was varied such that the sensitivity of both algorithms was comparable to that of colposcopy in expert hands. The specificity of the algorithms were then compared to that of colposcopy in expert hands. Table 3 indicates that the multivariate statistical algorithm has a similar sensitivity and significantly improved specificity for differentiating diseased from non diseased tissues relative to colposcopy in expert hands. A comparison between the algorithm based on the entire emission spectrum and that based on intensity at two emission wavelengths indicates that the latter algorithm performs with a similar sensitivity and specificity. Note that exclusion of samples with inflammation from the group of nondiseased tissues results in a 9% increase in the specificity of the multivariate statistical algorithm, applied to both the calibration and prediction sets.

DISCUSSION AND CONCLUSIONS

The primary goal of this study was to develop and test a method for constructing a diagnostic algorithm that intelligently extracts all the diagnostically useful spectral information and provides a clinically useful classification output. The objectives in developing this new algorithm were to: (1) preprocess spectral data for interpatient and intrapatient variation without the need for a priori information, (2) dimensionally reduced spectral data with minimal information loss, and (3) assign posterior probabilities to the diagnosis of each sample, such that samples with relative uncertain diagnosis can be identified for possible reevaluation by a clinician.

Preprocessing spectral data is a critical step in development of the algorithm; interpatient and intrapatient variation in the fluorescence intensity and line shape of a particular tissue type are frequently equal to or greater than variations between tissue categories. In this report, three methods of preprocessing are compared for their ability to reduce interpatient and intrapatient variation in the spectral features of a particular tissue type while maintaining inter-category differences. The three preprocessing techniques examined here all have the advantage over preprocessing methods used for the previous two-stage algorithm [12–14] in that colposcopic identification of a reference normal site prior to spectroscopic analysis is not required for algorithm implementation. Mean-scaling of original spectra serves to calibrate for interpatient variation in fluorescence intensity and line shape. However,

the large intrapatient variation of normal tissue fluorescence intensity in some patients (Fig. 2a) precludes accurate diagnosis using this method. Normalization of tissue spectra calibrates for the large interpatient and intrapatient variation in fluorescence intensity. However, the large interpatient variation in fluorescence line shape is not accounted for by this method of preprocessing. Mean-scaling normalized spectra from each patient calibrates for both the interpatient variation in fluorescence line shape and for the inter- and intrapatient variation in normal tissue fluorescence intensity. As a result, there is much less intrapatient variation in the resulting preprocessed spectra of a diagnostic category hence providing better discrimination. An added advantage of using normalized, mean-scaled spectra is that it does not require measurement of absolute intensity information.

The current clinical study utilized a single pixel probe to collect spectral information from a 1 mm diameter area of the cervix and fluorescence spectra were acquired from approximately two nondiseased sites and two diseased sites identified colposcopically in each patient. Spectra were then preprocessed for interpatient and intrapatient variation by mean-scaling normalized spectra from each patient. Mean-scaling calibrates for interpatient variation most effectively when there are approximately equal number of nondiseased and diseased spectra from each patient as was the case in the current study. However, in a different study design, the proportions of nondiseased and diseased spectra may no longer be approximately equal for each patient, and the method of mean-scaling may not be as effective as in the current study. Therefore, in addition to mean-scaling, other preprocessing methods will have to be tested for optimal algorithm development. Two potential methods are median scaling and maximum intensity scaling. In median scaling, the median spectrum is calculated from each patient, and substracted from each spectrum in that patient. Although this method requires the presence of at least one nondiseased and one diseased spectrum from each patient, it is likely that the ratio of the number of diseased to nondiseased spectra will not affect preprocessing in this manner significantly. Alternatively, all fluorescence spectra acquired from a patient could be normalized by the peak intensity of the spectrum with the maximum fluorescence intensity in that patient. This method of maximum intensity scaling is based on the premise that if a large number of sites are sampled in each patient, the one with the maximum fluorescence intensity can be expected to be histologically normal. This assumption is valid in general; our studies have shown that $< 1\%$ of patients have an abnormal lesion covering the entire cervix.

Compression of spectral data using PCA enables exploration of the diagnostic contribution of the entire fluorescence emission spectrum. For example, PCA of tissue spectra at 337 nm excitation demonstrates that two principal components that are uncorrelated with each other can describe $> 95\%$ of the variation in the original data set. Furthermore, component loadings can be used to select the diagnostically most important features of the spectra. Evaluation of the component loadings of normalized, mean-scaled spectra at 337 nm excitation demonstrates that intensity at only two emission wavelengths can be used to develop an algorithm that performs equally well as the algorithm that employed the entire emission spectrum. The reduction in the required number of emission intensities implies reduction in the cost of the clinical system used to measure the spectroscopic data. For example, if intensity is acquired at only two emission wavelengths, the polychromator and intensified diode array can be replaced by a motorized filter wheel (containing two bandpass filters) and a photomultiplier tube to detect the intensities only at the two selected emission wavelengths.

The calculation of posterior probabilities using logistic discrimination not only determines each sample's diagnostic category, but also the certainty of the classification. The availability of this information during diagnosis would allow the clinician to reassess sites that are classified with higher relative uncertainty. Furthermore, the boundaries dividing diagnostic categories can be chosen to minimize either false-negatives or false-positives (adjusting sensitivity and specificity) (Fig. 7) as required by a particular clinical setting. In addition, since the prevalence of the disease in the form of prior probabilities was used to construct the decision process, logistic discrimination can be used to adjust the decision surface to fit different patient populations.

Several groups have developed algorithms using multivariate analysis to extract diagnostic information from tissue fluorescence spectra [19,20]. Schomacker et al. [19] used multivariate linear regression to develop scores for adenamatous and hyperplastic colonic polyps and then used a binary decision line for classification.

O'Brien et al. [20] used a variety of multivariate techniques including principal component analysis and multivariate linear regression to develop classification scores for atherosclerotic and normal aorta. Again, a binary decision line was used to discriminate between these two tissue types [20]. The multivariate statistical algorithm that our group has developed first preprocesses spectral data for interpatient variation, dimensionally reduces the data into an orthogonal set of principal component scores, which are then entered into a classification algorithm based on logistic discrimination that assigns posterior probabilities to the diagnosis of each sample. Furthermore, the component loadings of PCA are used to relate the principal component scores of the samples to their original fluorescence emission spectra. The utility of the component loadings for reducing the number of emission variables required to achieve classification with a minimum decrease in predictive ability is also explored. There are two primary differences between the multivariate statistical algorithm that we have developed and those developed by other groups [19,20]. We use logistic discrimination for classification rather than a linear decision line. The advantage of using logistic discrimination rather than a binary decision is that a probability-based classification is clinically more versatile as it assigns a relative uncertainty to samples that are close to decision boundaries. Furthermore, although both Schomacker et al. [19] and O'Brien et al. [20] used multivariate statistical techniques to develop diagnostic scores, neither of these groups attempted to relate the scores that were developed to the original variables of the tissue fluorescence spectra, which is critical to the understanding of the basis of the diagnostic algorithm.

Another approach to extracting the diagnostic information from tissue fluorescence spectra is to model the physical basis of the measured signal [21]. This allows not only extraction of diagnostic information but also provides insight into the physical basis of the measured spectra. The advantage of using a physically based model [21] over a statistically based model (PCA) [22] is that the former method provides insight into the biochemical and morphologic basis of the diagnostic algorithm. However, there are two primary advantages to using a statistical model such as PCA. Although a linear method of analysis, PCA can still be used effectively to model the nonlinear turbid tissue fluorescence because it does not directly model the physical basis of the fluorescence. Because it is less restrictive, PCA can permit a better fit to the real data than physical models can, through the use of a linear combination of orthogonal components [22]. In addition, the principal components are uncorrelated to each other and therefore can be used in a variety of classification algorithms, which generally require uncorrelated variables, and each principal component can be correlated to each of the original emission variables, thus providing insight into components of the original fluorescence emission spectrum that are diagnostically useful.

Evaluation of the multivariate statistical algorithm at 337 nm excitation indicates that whereas normal squamous tissues can be differentiated from SILs with a high sensitivity and specificity, normal columnar epithelia and inflammation are indistinguishable from SILs. Furthermore, this algorithm cannot differentiate high grade from low grade SILs effectively. These limitations must be solved before the technique can be implemented clinically. Analysis of in vivo cervical tissue fluorescence spectra at 380 nm and 460 nm excitation indicates that spectra at these wavelengths can overcome the diagnostic limitations at 337 nm excitation. Spectra at 380 nm excitation can be used to differentiate diseased tissues from normal columnar epithelia and inflammation with an average sensitivity and specificity of 77%±1 and 72%±9, respectively; spectra at 460 nm excitation can be used to differentiate high grade SILs from low grade SILs with an average sensitivity and specificity of 80%±4 and 76%±5, respectively. These results described in a companion work [15] indicate that spectra at multiple excitation wavelengths are essential for the differential diagnosis of SILs at colposcopy.

REFERENCES

1. Wright TC, Kurman RJ, Ferenczy A. Cervical Intraepithelial Neoplasia. In: Blaustein A, ed. "Pathology of the Female Genital Tract." New York: Springer-Verlag, 1994, pp 156–177.
2. Cancer Facts and Figures. American Cancer Society, 1994.
3. Koss LG. The Papanicolau test for cervical cancer detection. JAMA 1989; 261:737–743.
4. Mitchell MF. Preinvasive diseases of the female low genital tract. In: Greshenson DM, DeCherny A, Curry S, eds. "Operative Gynecology." Philadelphia: W.B. Saunders. 1993, pp 231–256.
5. Mitchell MF. Accuracy of Colposcopy. Consultations in Obstetrics and Gynecology 1994; 6(1):70–73.
6. Cothren RM, Richards-Kortum RR, Rava RP, Boyce GA,

Doxtader M, Blackman R, Ivanc T, Hayes GB, Feld MS, Petras RE. Gastrointestinal tissue diagnosis by laser-induced fluorescence spectroscopy at endoscopy. Gastrointest Endosc 1991; 36:105–111.

7. Schomacker KT, Frisoli JK, Compton CC, Flotte TJ, Richter JM, Nishioka NS, Deutsch TF. Ultraviolet laser induced fluorescence of colonic tissue: Basic biology and diagnostic potential. Lasers Surg Med 1992; 12:63–78.
8. Lam S, Hung JYC, Kennedy SM, Leriche JC, Vedal R, Nelems B, Macaulay CE, Palcic B. Title Am Rev Dis 1992; 146:1458–1461.
9. Glassman WS, Liu CH, Tang GC, Lubicz S, Alfano RR. Ultraviolet excited fluorescence spectra from non-malignant and malignant tissues of the gynecologic tract. Lasers Life Sci 1992; 5:49–58.
10. Lohmann W, Muβmann J, Lohmann C, Kunzel W. Fluorescence of the cervix uteri as a marker of dysplasia and invasive carcinoma. Eur J Obstet Gynecol Reprod Bio 1992; 31:249–253.
11. Mahadevan A, Mitchell MF, Silva E, Thomsen S, Richards-Kortum RR. Study of the fluorescence properties of normal and neoplastic human cervical tissues. Lasers Surg Med 1993; 13:647–655.
12. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Silva E, Richards-Kortum RR. Fluorescence spectroscopy: A diagnostic tool for cervical intraepithelial neoplasia (CIN). Gynecol Oncol 1994; 52:31–38.
13. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Richards-Kortum RR. In vivo diagnosis of cervical intraepithelial neoplasia (CIN) using 337 nm laser induced fluorescence. Proceedings of the National Academy of Sciences 1994; 91:10193–10197.
14. Ramanujam N, Mahadevan A, Mitchell MF, Thomsen S, Silva E, Richards-Kortum RR. Fluorescence spectroscopy of the cervix. Clinical Consultations in Obstetrics and Gynecology 1994; 6(1):62–69.
15. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Malpica A, Wright T, Atkinson N, Richards-Kortum RR. Spectroscopic diagnosis of cervical squamous intraepithelial neoplasia in vivo using laser induced fluorescence spectra at multiple excitation wavelengths. Lasers Surg Med 1995 (in press).
16. Dillon RW, Goldstein M. "Multivariate Analysis: Methods and Applications." New York: John Wiley & Sons. 1984.
17. Devore JL. "Probability and Statistics for Engineering and the Science." Pacific Grove, CA: Brooks/Cole, 1992.
18. Albert A, Harris EK. "Multivariate Interpretation of Clinical Laboratory Data." New York: Marcel Dekker, 1987.
19. Schomacker KT, Frisoli JK, Compton CC, Flotte TJ, Richter JM, Deutsch TF, Nishioka NS. Ultraviolet laser-induced fluorescence of colonic polyps. Gastroenterology, 1992; 102:1155–1160.
20. O'Brien KM, Gmitro AF, Gindi GR, Stetz ML, Cutruzzola FW, Laifer LI, Deckelbaum LI. Development and evaluation of classification algorithms for fluorescence guided laser angioplasty. IEEE Trans Bio Eng 1989; 36(4):424–431.
21. Richards-Kortum RR, Rava RP, Fitzmaurice M, Kramer JR, Feld MS. 476 nm excited laser-induced fluorescence spectroscopy of human coronary arteries: Applications in cardiology. Am Heart J 1991; 122(4):1141–1150.
22. Haaland DM. Multivariate calibration methods applied to quantitative analysis of infrared spectra. In: Jurs PC, ed. "Computer-Enhanced Analytical Spectroscopy," Vol. 3. New York: Plenum Press, 1992.